DOCUMENT RESUME

ED 334 216 TM 016 668

AUTHOR Balke, Gudrun
TITLE Multilevel Factor Analysis of Proficiency in English as a Foreign Language.
PUB DATE Apr 91
NOTE 15p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL, April 3-7, 1991).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Achievement Tests; *Elementary School Students; *English (Second Language); *Factor Analysis; Foreign Countries; Grade 5; Intermediate Grades; *Language Proficiency; Language Tests; *Second Language Learning
IDENTIFIERS LISREL Computer Program; *Multilevel Analysis; *Sweden; Variance (Statistical)

ABSTRACT

As part of the National Evaluation Program of the Swedish school system, proficiency in the English language is evaluated for students in grade 5 (at age 12 years). Variation between classes and variation within classes were studied in a multilevel factor analysis (MFA) using a sample of 3,162 students in 162 classes for whom student questionnaire (self-assessment), teacher questionnaire, and achievement test results were available. In all, eight achievement tests were administered, which evaluated reading comprehension and vocabulary knowledge, listening comprehension, and writing proficiency. Only a small subsample took a speaking test. A two-level MFA was performed using the LISREL VI computer program. Analyses from the student questionnaire and tests revealed factor structures for: (1) proficiency as measured by tests; (2) proficiency as measured by self-assessments; (3) attitudes concerning learning English and using English; and (4) frequency of leisure-time use of English. The factor structure based on total variance was similar to that of the within-class part. The MFA provided additional information about the effects of group membership and instructional effectiveness. Five tables present study data. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED334216

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

GUDRUN BALKE

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

MULTILEVEL FACTOR ANALYSIS OF PROFICIENCY IN ENGLISH AS A FOREIGN LANGUAGE

Gudrun Balke
University of Gothenburg, Sweden

Paper presented at the symposium 'Multilevel Factor Analysis: Applications to Education' at the annual meeting of the American Educational Research Association, Chicago, April 3-7, 1991.

Gudrun Balke
Dept of Education and Educational Research
University of Gothenburg
Box 1010
S-431 26 Mölndal
Sweden

TM016468

BEST COPY AVAILABLE

INTRODUCTION

The National Evaluation Program of the Swedish school system.

The National Evaluation Program comprises a vast amount of different tasks which are to be solved by using different methods on different aggregation levels. The point of departure is the present curriculum and the aim of the evaluation is to get a view, as comprehensive as possible, of the school system.

Included in the National Evaluation Program is a recurrent assessment of pupils' knowledge, skills and attitudes. The first round of this assessment was conducted in the spring of 1989. It comprised assessment of pupils in grade 2 (9 years old) and in grade 5 (12 years) in the compulsory school. Future phases will comprise assessments of students at higher levels in the school system.

The data collected cover all school subjects, except for physical education, taught in the grades mentioned above. The collection is made in a 3 per cent sample of classes, stratified according to size of schools. All classes (pupils) in the samples were assessed in all school subjects. The data collection was administered by the teachers during the school days.

Evaluation of English in the compulsory school.

In the present paper the evaluation of English in grade 5 is presented. The teaching of English starts in grade 4, which means that the assessment is made after not quite two years of study. Three lessons (40 minutes each) were spent on the data collection (tests and questionnaire). Beside this, a speaking test was administered to a subsample outside this time limit.

The goals for English teaching emphasized in the curriculum from 1980 are based on the concept of communicative competence. The pupils are to be trained to handle a functional language for communication in authentic settings. The spoken language is to dominate the teaching and it is important to encourage the pupils to use their own English even if mistakes are made. These goals are quite different from those expressed in the two previous curricula, published in 1961 and in 1968, where teaching should be aimed more at linguistic elements and grammar.

Purpose

The descriptive outcomes of the evaluation have been presented in different reports (in Swedish), and in further studies relations between different factors are analysed both at individual level and at class (teacher) level.

What is not clear in these analyses is how much of the total variance can be referred to as variation between classes and how much can be referred to as variation within classes. In the study to be presented here the pupil variance will be divided into within class and between class variation. This is carried out in a Multilevel Factor Analysis, presented in Muthén (1990).

METHOD

Sample

The sample in grade 5 consists of around 3400 pupils in 164 classes. At this stage (i e the middle stage: grade 4 to 6) of the comprehensive school the teachers teach, with some individual exceptions, every subject. This means that there were about 164 class teachers involved in the evaluation. As far as the evaluation of English is concerned, information was received from 3438 pupils in 163 classes. Complete data, i e results on both the pupil questionnaire and of all the tests, are available for 3162 pupils in 162 classes. The size of the classes varies from 3 up to 30. The small classes are taught together with classes in other grades.

Data

Teacher questionnaire.

In order to get information about how the goals expressed in the curriculum are implemented in the classroom, a questionnaire was given to the teachers. In this, many different aspects of the way teachers think they teach are asked for. A further intention is to find out why, in certain cases, the teaching is not in line with the one stated in the curriculum. Thus, the teachers are also asked to assess their knowledge and skill in English and their teaching competence. Questions where the teachers are asked to evaluate their teacher education and training are also included.

Achievement tests.

In the 8 achievement tests given to all pupils, reading comprehension and vocabulary knowledge, listening comprehension and writing proficiency are tested. A test of speaking proficiency was, for practical reasons, administered by personnel from the project. Because of this, speaking proficiency could only be tested in a subsample.

The achievement tests given to the total sample and included in the analysis to be presented here are:

Posters: 9 different written posters are presented. The task is to relate each with one out of 11 given sentences. This task tests reading comprehension.

Crossword. 8 written sentences in English are given. In each sentence one word is missing. The task is to find the missing word and write it in the appropriate place. When all missing words are found, a new one is given vertically. The clue for this word is written underneath. This task tests reading comprehension and vocabulary.

Gaps: 10 very simple sentences are written in English. In each sentence one word is missing and what word it is is very evident. This task tests primarily vocabulary knowledge.

Gaps, multiple choice: 10 written sentences, each with one gap, are given. The word to be put in the gap is given together with three distractors. This task tests vocabulary and reading comprehension.

Cloze: A coherent text (a letter) with 12 gaps is given. The task is to find out what the missing words are from the context only. This type of test is frequently used in language testing and is supposed to test an integrative language skill.

Dialogues: 6 short dialogues are given on a tape. The task is to find out the topic of each dialogue. The pupil is to choose the correct topic among 4 given in the test form. This multiple choice task tests listening comprehension.

Note taking: 7 short dialogues are presented on a tape. In each dialogue a name is spelled or figures (e g a telephone number) is given. The task is to note the name or the figures. This task tests a limited part of listening comprehension.

Letter. A letter from an English boy or girl ('Chris') is given in the test form. The task is to write a reply letter. This is a writing proficiency test.

The tests aim at defining the level of achievement to be expected after two years of study. As described above, half of the tests of the receptive skills are mainly of the kind where pupils do the tasks by writing single words, while the other tests are of a non-productive (mostly multiple choice) type. In the scoring of the productive tests spelling mistakes are accepted as long as the word can be assumed to be understood by a native speaker.

In the writing test ('Letter') the pupil produces a text of his own and in the speaking test he is put into three different situations in which he has to act orally. These two tests of the productive skills are scored on a scale with four levels. The definitions of the levels are based on communicative aspects. The scoring of the writing test was made by the teachers.

Pupil questionnaire.

Since oral proficiency, risk-taking and willingness to use English are the most important goals and since these aspects are best, but not fully, captured by the speaking test (which was only given to a subsample), supplemental information was collected from all participating pupils. Several self-assessment tasks, mainly aimed at the ability to handle the spoken language, are thus included in the questionnaire given to the pupils. Some of these tasks are strictly self-assessment ones while others are more directed towards attitude to and motivation for learning and using English.

When teaching is evaluated by assessing pupils' knowledge, skills and attitudes, it must be considered what may be an effect of the teaching and work in school and what is learnt in life outside school. In Sweden, most people (especially youngsters) hear and use English in varying degrees every day. It is thus as important to find out type and frequency of leisure time activities 'in English' and to relate this information to the test and self-assessment results, as it is to relate information about the teaching and school work to the results. Questions about the pupils' activities are thus included in the questionnaire.

The pupil questionnaire comprised 31 questions and tasks. Some of these are divided into subquestions, which gives a total of 40 questions and assessment tasks. Of these, 13 are self assessments, 5 are attitude assessments and 12 are questions concerning leisure time activities. The remaining questions are mainly directed toward previous expectations of learning English and toward school activities.

Analyses

As mentioned above, the general findings have been reported elsewhere and they are summarized below. Beside that, the data have been analysed descriptively from specific aspects. One is sex differences in test results and in responses to the questionnaire, and another is characteristics of pupils with low results on all tests.

Different factor analyses, based on the pupil and the class level, have also been performed. These analyses revealed stable structures on both levels. Thus, the analyses of the data from the pupil questionnaire and tests revealed factor structures for proficiency as measured by the tests, proficiency as measured by the self assessment tasks, attitudes of learning and using English and the frequency of leisure time activity (in English). The analyses of the factor structures made on the pupil and the class levels are to be reported elsewhere. The present paper will focus on the Multilevel Factor Analysis as presented by Muthén (1990).

A two-level MFA of the pupil data was performed by LISREL VI (Jöreskog and Sörbom, 1986). The analysis is based on the covariance matrix of the class means (S_b) weighted by class size, and the pooled-within covariance matrix (S_{pw}), as explained in Muthén (ibid).

In order to get a manageable model to be analysed by the MFA, the number of variables to be included has to be limited. The selection of variables was made from previous LISREL analyses and those included in the MFA are those which were shown to be stable and highly loading.

RESULTS

Summary of general findings

Practically all teachers are formally qualified to teach English. Around 90 per cent graduated before the present curriculum took effect. In spite of this, very few have had in-service training in accordance with the (new) goals specified in the present curriculum. Consequently, more than half of the teachers state that they need more knowledge of foreign language learning and teaching methodology and think they need to improve their own proficiency, especially the oral one.

Regarding teaching content, organization of the teaching, use of textbooks, activities in the classroom, homework and classroom tests, a very clear picture comes out. What we see is a very traditional teaching, not at all congruent with the goals stressed in the curriculum. All responses about the actual teaching are more in accordance with the previous curricula (from 1962 and 1969) than with the present one.

In spite of this, all participating teachers find the emphasis in the present curriculum on communicative competence justified. The discrepancy between the actual teaching on the one hand and a teaching based on the goals stated in the curriculum on the other hand can, at least to a certain degree, be explained by the inadequate competence of many of the teachers. The more uncertain the teachers are of their own competence, the more traditional is their teaching and the more dependent on textbooks they are.

The pupils are often exposed to English outside school. A great majority listen to English on TV and to English songs almost every day. Other rather frequent activities are listening to English radio programs and playing computer and role games. The pupils' answers to the questions about the activities in school show the same picture as the teachers' answers, mentioned above. A finding in line with this is also that the pupils are given few opportunities to speak English during the lessons.

The pupils' self assessments refer, as mentioned above, mainly to their use of spoken English. When the responses are checked in the subsample that took the speaking test, it is shown that the self assessments are valid measurements of oral proficiency and of self assurance.

Generally, the pupils have positive attitudes towards the use of English and are also rather confident in their use of English in real settings, though most of them state that they are able to say only a few words. Concerning English as a school subject, the picture is somewhat different. A majority of the pupils think that it is fun to learn English in school and to speak English during the lessons, but compared to the attitudes towards English outside school, the enthusiasm about English as a school subject is lower.

Regarding the results on the achievement tests, it is found that ninety per cent of the pupils attain a basic level of proficiency in each of the four skills. A majority of the 10 per cent who

fail in one skill also fail in the others. Most of the tests of the receptive skills are rather easy, which causes some ceiling effects. In the tests of the productive skills the pupils can better show their full communicative ability.

A salient result is the great variation between classes both in test results and in the self assessments of school performance.

Analyses of the structure of achievement, self assessment and leisure time activity

As mentioned above, the selection of variables to be included in the MFA was made from previous LISREL analyses. Those included were shown to be stable and highly loading. These variables are presented in Table 1 together with the means, standard deviations and ranges.

Insert Table 1 about here

The 15 questionnaire items included are directed towards self assessment of proficiency and leisure time activity. The answers on these items are given on scales with 4 or 5 fixed alternatives. All the 8 achievement tests, taken by everyone in the sample, are also included in the analysis.

The first two self assessment tasks shown in Table 1 are formulated in the same way. They capture the pupil's opinion of how well he/she can handle the spoken language and were summed into one variable. The third task in Table 1 is a judgement of how well the pupil will manage when being together with English-speaking persons. The pupil is also asked to estimate how much he/she understands when the teacher speaks English. Thus, the first two assessments (summed into one) concern general (listening and speaking) competence, the third refers to competence in a social context and the last one is an assessment of listening competence in school. All these variables refer to the use of spoken English.

The question concerning the frequency of leisure time activity is divided into 10 defined activities and the pupils estimate their degree of activity on a frequency scale with four steps. The activities are summed on the basis of what skill is activated: listening, reading, writing and speaking. Since only one writing activity is included (writing letters), this is put together with the reading activities. This gives three variables capturing activity 'in English': listening, reading/writing and speaking.

The analysis of the tests is to be presented elsewhere. It reveals, however, that only one single factor is captured by the tests. Thus, no special factor of, for instance, listening comprehension, can be defined. This may be explained by the fact that the testing is made at an initial stage of learning, i e after less than two years of study.

Since the construction of the tests was based on the four different skills, the tests are, before the MFA, summed according to what skill is supposed to be tested. Thus, the tests based on reading comprehension and vocabulary knowledge are summed into one variable and the same is done with the two listening comprehension tests. Writing proficiency is tested by one test only. Since the cloze test is looked upon as an integrative test, this is kept as a separate variable.

The summing of some items and tests gives variables which differ in scale and range. In order to make variables technically more manageable, averages are used instead of sums.

The final set of variables and their means, standard deviations, ranges and intra-class correlations ($r_{(i-c)}$) are given in Table 2.

Insert Table 2 about here

Of the 10 variables, three are self assessments, three are estimates of frequency of leisure time activity and four are test results. The $r_{(i-c)}$'s are low, indicating small variance components at the between level compared to the total variances. It may be noted, however, that the intra-class correlations are higher for the tests than for the other variables.

Factor analysis based on total variance

In Table 3 the result of the factor analysis based on the individual (total) results is presented.

Insert Table 3 about here.

As mentioned before, only one language (tested) proficiency factor could be established. Beside this factor, two more are identified: 'self assessed proficiency' and 'language use'.

The primary purpose of the self assessment tasks was to supplement the information about proficiency in English given by the tests. This means that the self assessments should partly overlap with the tests. Consequently, the tests and the self assessments were put together into one factor, 'language proficiency'. This factor is strongly related to the tests of reading comprehension and vocabulary knowledge, the cloze test, and, somewhat less, to the listening tests. The lowest loading for a test is observed for the writing proficiency test. Concerning the self assessments, the one of school (listening) competence has a somewhat higher loading on 'language proficiency' compared to the other two self assessments. The lowest loading on this factor is the one on the assessment of a general competence of spoken English.

The second factor defined in the model is 'self assessed proficiency'. This factor is related to the self assessments when language proficiency is partialed out. As mentioned previously, one of the purposes with the self assessment items was to get information about communicative competence, oral proficiency and willingness to use English, i e those aspects of language learning emphasized in the present curriculum, but not captured by the 8 tests. Looking at the results presented in Table 3, this aim seems to be fulfilled. The self assessments partly measure assessments of capabilities also measured by the tests, but, in addition, the residuals form a factor unrelated to language proficiency. This factor is loaded by the self assessment of general competence and somewhat less by the assessment of sociolinguistic competence. The loading of the assessment of school (listening) competence is the lowest one on this residual factor.

The 'self assessed proficiency' factor may be looked upon as a self esteem factor when all sorts of competence are partialed out. However, its relation with the 'language use' factor indicates that it comprises more than just pure self esteem. Furthermore, in previous analyses, this 'self assessed proficiency' factor is shown to be related to attitudes of learning and using English as well as to the result of the speaking test, given to a subsample. These facts support the conception of this factor as a factor measuring communicative competence as well.

The third factor is 'language use'. The highest loading on this factor is obtained for the reading and writing activities. The loadings for the two other activity variables are about equally high.

The 'language use' factor has a rather high correlation with 'self assessed proficiency' (.529) and a moderate one with the 'language proficiency' factor (.229). The chi-square of this model is 120.10 (df 30) and the goodness of fit index is .992 (.986).

Multilevel Factor Analysis

The analysis in Table 3 is based on total variance between individuals. In this analysis, the data are treated as if sampled at the individual level, when the sampling unit in reality is 'class'. In order to make an analysis at both levels simultaneously, a two-level MFA was performed using the LISREL VI program (Jöreskog and Sörbom, 1986).

The result is shown in Table 4. The loadings are, as in Table 3, standardized according to the standard deviations of the variables. The standardization is made from the estimated variances given in the MFA model.

Insert Table 4 about here.

The model fits the data well. Chi-square is 182.79 (df 69) and the goodness of fit index for within is .991 and for between .944. The chi-square value is in fact a quasi-chi-square value given by the MUML estimator (Muthén, 1990). This is, however, likely to be close to the one given by the true ML procedure (Muthén, ibid.).

Three factors are defined at the within level. They are the same as those found in the analysis based on total variance: one 'language proficiency' factor, one 'self assessed proficiency' factor and one 'language use' factor. The loadings at the within level are, with one exception, also similar to those in the previous analysis. The exception is the loading for the writing test which is higher at the within level.

At the between level, only one factor emerges. This is loaded by the self assessment and test variables. The loadings of the leisure time activity variables are insignificant and not included. The factor at the between level is thus, as far as the variables are concerned, parallel to the 'language proficiency' factor at the within level.

At the within level, 'language proficiency' is dominated by the reading/vocabulary tests and, secondly, by the cloze test . The listening tests and the writing test have lower loadings. The self assessment of school (listening) competence has, as in the previous analysis, the highest loading of the self assessments. The variables with the lowest loadings on this factor at the within level are the self assessments of sociolinguistic competence and general (listening and speaking) competence.

At the between level, the 'language proficiency' factor is dominated by the same two tests as at the within level: the reading/vocabulary tests and the cloze test. The self assessment of socio-linguistic competence is the third variable of this factor at the between level. The loading of this variable is even somewhat higher than the one for the listening comprehension tests. The self assessment of school (listening) competence comes next in size and the self assessment of general (listening and speaking) competence is, as at the within level, the least important self assessment variable. At the between level the test of writing proficiency has an insignificant loading.

The other two factors at the within level are 'self assessed proficiency' and 'language use'. As said before, the loadings are similar to those obtained in the analysis based on total variance (Table 3). Thus, the self assessed general competence dominates 'self assessed proficiency',

followed by the assessment of sociolinguistic competence. The least important variable is the assessment of school (listening) competence. The 'language use' loading for reading and writing activities is the highest one in this factor. The loadings of the two other activity variables, listening and speaking, are of about the same size.

The correlations between factors are also similar to those found in the previous analysis. The correlation between 'language use' and 'self assessed proficiency' is rather high, .526, and the one between 'language use' and 'language proficiency' rather low, .237.

In order to get information about the proportions of variance explained in the MFA, the result given in Table 4 has been transformed to components of variance. These are shown in Table 5.

Insert Table 5 about here

The decompostion of the variance is based on the estimates of variance given in the MFA model. These estimates are rather close to the sample variances, as indicated by the high goodness of fit indices presented in Table 4.

The first two pairs of columns in Table 5 show the proportions of within and of between variance respectively. In the four remaining columns the proportions of total variance are given. From these four columns it can be seen that the variance components at the between level are much smaller than those at the within level. This is especially valid for the self assessments. The between components of the test variables are somewhat higher than 10 per cent. This is about the same result as shown by the intra-class correlations in Table 2. The $r_{(i-c)}$'s are, however, based on sample variances and not, as the percentages in Table 5, on the variances estimated in the MFA model. The similarity between the two types of measurement indicates that the model fits the data well at the between level, which is also given by the goodness of fit index of .944 in Table 4.

At the within level, the explained variance is greater than the unexplained variance for the reading/ vocabulary tests and for the cloze test. The listening tests and all three self assessments have about the same proportions of explained and unexplained variance. A greater proportion unexplained variance is shown in the writing test and the reading/writing activity. For the two other activities only 20 per cent of within variance is explained.

At the between level, almost all variance of the reading/vocabulary tests is explained. The cloze test, the self assessment of sociolinguistic competence and the listening tests have greater explained components than unexplained ones. For the assessments of school and general competence the explained variances are 33 and 17 per cent of the estimated between variances. The lowest component of explained variance at the between level is that of the writing test - only 2 per cent of the estimated between variance and .3 per cent of the estimated total variance.

DISCUSSION

The factor structure given in the analysis based on total variance is similar to the structure shown in the within part of the MFA analysis. This is due to the rather low proportions of between variance, shown both in the $r_{(i-c)}$'s in Table 2 and in the percentages in Table 5. This is, in turn, due to the type of variables included, i e variables measuring general performance. Thus, a similarity between total and within might also show up in other studies of the same kind as the present one. The Multilevel Factor Analysis provides additional information illuminating important effects of group membership.

Consider the result of the writing test. This test is rather important at the within level, but insignificant at the between level. This result is probably an effect of the scoring of this test. As mentioned earlier, the pupils' products were scored by their teacher on a scale with four steps. Formulated criteria and examples of products at each step were given in the instructions to the teachers. It seems as, if each teacher has been able to differentiate between his/her pupils, but the teacher cannot find the correct position of the class on the scale, in spite of careful directions. This indicates that comparisons of different classes should be made above teacher/ class level.

The assessment of sociolinguistic competence is a strong variable of the 'language proficiency' factor at the between level.The loading of this assessment is even close to that of the cloze test. In the assessment of sociolinguistic competence the pupil is asked how well he/she will manage when being together with English-speaking persons. The high loading of this assessment at the between level might be a result of different experiences of meeting English-speaking persons, and thus dependent on regional and social settings, as well as a result of the teacher's willingness and proficiency of speaking English during the lessons. In both the teacher and pupil questionnaires it is shown that the teachers differ very much in how often they speak English.

The self assessment of school (listening) competence refers directly to the teacher's speech. This assessment of how much the pupil understands his/her teacher is the most powerful assessment of 'language proficiency' at the within level. Thus, the strongest self assessment in 'language proficiency' between classes is related to how often the pupils (class) are exposed to English speech, while the strongest assessment within classes is related to how much the pupil understands this speech.

The self assessment of a general listening and speaking competence, doesn't refer to any special setting. This assessment is the weakest one in the factor 'language proficiency' at both levels. The within factor 'self assessed proficiency', unrelated to the 'language proficiency' factor, has, however, the highest relation to the assessment of general competence.

'Language proficiency' has higher loadings on the reading/vocabulary tests, the assessment of sociolinguistic competence and the listening tests at the between level compared to the within. This might be explained by practice and teaching, i e how used the pupils (class) are in reading different texts, in acting in oral settings and in listening to the teacher's speech. This explanation should also comprise the writing test, but it doesn't do that, since the scoring of this test didn't work on the class' level.

Reading comprehension and vocabulary knowledge seem to be highly related to 'language proficiency' at both levels, i e both to each pupil's achievement and to the achievement level of the class. It should be remembered that the tests measuring these two types of achievement are many and varying compared to the other types of achievement (listening and writing). Thus, the importance of these tests may have been over-emphasized in this study.

Beside the language proficiency factor, two other factors emerge at the within level. The leisure time activities given in the questionnaire are mostly available to all and thus not dependent on regional or social setting, which in turn can be related to school or class. Confidence in using English for communication, when performance related to tested proficiency is disregarded, is also dependent on the individual and not the class to which he belongs, especially when the teaching is as it is: traditional and emphasizing formal knowledge.

'Self assessed proficiency' comprises assessments of proficiency not captured by the tests.The proficiency assessed is aimed at the handling of spoken English for communication. Remembering the traditional teaching, it is not surprising that the correlation between 'language use' and 'self assessed proficiency' is high, much higher than the correlation between 'language use' and 'language proficiency'. This is quite in line with the previous results about the teaching as traditional and based on textbooks and not the pupils' own English.

The next step in the MFA will include variables at class level, such as teachers' competence and teaching method. This is expected to give new and interesting information about relations between pupil proficiency and teacher competence and teaching.

REFERENCES

Jöreskog, K-G & Sörbom, D. (1986). LISREL VI. User's guide. Mooresville: Scientific Software Inc.

Muthén, B. O. (1990). Multilevel factor analysis of class and student achievement components. UCLA Statistics Series #76.

Table 1. Means, standard deviations and ranges of the original questionnaire items and tests. N=3162.

Questionnaire item/Test	mean	sd	range
Self assessment:			
of understanding spoken English, general	4.30	.54	1-5
of speaking English, general	4.16	.54	1-5
of managing social intercourse 'in English'	3.79	.86	1-5
of comprehending the teacher speaking English	2.98	.65	1-4
Leisure time (English) activity:			
radio	1.96	1.18	1-4
TV (subtitled)	3.28	.76	1-4
TV (not subtitled)	1.89	1.15	1-4
music	3.34	.96	1-4
cartoons	1.31	.68	1-4
newspapers, books	1.24	.60	1-4
games (computer, role)	2.15	1.16	1-4
pen-friends	1.33	.64	1-4
meets English speakers	1.17	.57	1-4
speaks English in the family	1.50	.88	1-4
Test:			
reading: posters (non-productive)	6.86	2.33	0-9
reading and voc.: crossword (productive)	5.17	2.33	0-8
reading and voc.: gaps (productive)	5.80	2.17	0-10
reading and voc.: gaps (non-productive)	7.31	2.02	0-10
cloze: letter (productive)	6.80	2.67	0-12
listening: dialogues (non-productive)	4.88	1.06	0-6
listening: taking notes (productive)	4.64	1.40	0-7
writing: letter	2.99	.90	1-4

Table 2. Means, standard deviations, ranges and intra-class correlations of the 10 variables. N=3162

Variable	mean	sd	range	$r_{(i-c)}$
Self assessment:				
general (listening and speaking) competence	4.23	.48	2-5	.038
sociolinguistic competence	3.79	.86	1-5	.028
school (listening) competence	2.98	.65	1-4	.040
Leisure time activities:				
listening	2.62	.61	1-4	.120
reading and writing	1.51	.48	1-4	.039
speaking	1.33	.57	1-4	.052
Tests:				
reading comprehension and vocabulary	2.51	.71	.1-3.7	.125
cloze	2.27	.89	0-4	.100
listening comprehension	1.90	.41	.2-2.6	.111
writing proficiency	2.99	.90	1-4	.136

Table 3. Factor analysis of test proficiency, self assessment and frequency of leisure time activity. Individual level, standardized solution. N=3162

	Language proficiency	Self assessed proficiency	Language use
Self assessment:			
general (list./speak.) competence	.384	.584	
sociolinguistic competence	.412	.516	
school (listening) competence	.482	.435	
Leisure time activity:			
listening			.483
reading/writing			.596
speaking			.443
Tests:			
reading compr./vocabulary	.916		
cloze	.822		
listening comprehension	.712		
writing proficiency	.561		
Language proficiency	-		
Self assessed proficiency	0	-	
Language use	.229	.529	-

Chi-square=120.10; df 30; goodness of fit=.992; adj goodness of fit=.986.

Table 4. MFA model (standardized) at within and between levels. N= 3162, G=162.

	Language proficiency		Self ass. prof.	Language use
	B	W	W	W
Self assessment:				
general (list./speak.) competence	.399	.390	.578	
sociolinguistic competence	.775	.405	.521	
school (listening) competence	.579	.480	.438	
Leisure time:				
listening				.487
reading/writing				.607
speaking				.436
Test:				
reading compr./vocabulary	.958	.911		
cloze	.855	.817		
listening comprehension	.757	.710		
writing proficiency	.156[a]	.620		
Language proficiency		-		
Self assessed proficiency		0	-	
Language use		.237	.526	-

Chi-square=182.79; df 69; Goodness of fit=.991/.944

a) Insignificant at 5 % level

Table 5. Decomposition of variances estimated in the MFA model in Table 4. Percentages.

	Parts of estimited Σ_{PW} / Σ_B				Parts of estimated Σ_T			
	within level		between level		within level		between level	
	expl	error	expl	error	expl	error	expl	error
Self assessment:								
general (list/speak) comp	48	52	17	83	47	50	.4	2
sociolinguistic comp.	43	57	60	40	43	56	1	1
school (listening) comp.	43	57	33	67	41	55	1	2
Leisure time:								
listening	24	76	-	100	22	67	-	11
reading/writing	37	63	-	100	36	61	-	3
speaking	21	79	-	100	20	76	-	4
Test:								
reading compr./voc.	83	17	92	8	73	15	11	1
cloze	67	33	73	27	60	30	7	3
listening compr.	50	50	57	42	44	44	6	5
writing proficiency	39	61	2	97	33	53	.3	13